$$\left\langle\, h_1,\, e_3,\ \left\{\begin{array}{l} h_4\text{:\_every\_q}\langle 0\!:\!5\rangle(x_6,\, h_7,\, \_),\ h_8\text{:\_linguist\_n\_1}\langle 6\!:\!14\rangle(x_6), \\ h_2\text{:\_have\_v\_1}\langle 15\!:\!18\rangle(e_3,\, x_6,\, x_9), \\ h_{10}\text{:\_a\_q}\langle 19\!:\!21\rangle(x_9,\, h_{12},\, \_),\ h_{13}\text{:\_obsession\_n\_1}\langle 22\!:\!32\rangle(x_9) \end{array}\right|,\ \{\, h_1 =_q h_2,\, h_7 =_q h_8,\, h_{12} =_q h_{13} \,\}\right\rangle$$

Taking (Out) Scope

(Recovery from Post-Sabbatical Syndrome)

Stephan Oepen

oe@ifi.uio.no

University of Oslo, Department of Informatics

SynSem Planning; February 8, 2017

Keeping me Awake Some Nights

My Introduction to the World of Dependencies

Who Did What to Whom?
A Contrastive Study of Syntacto-Semantic Dependencies

Angelina Ivanova[♣], **Stephan Oepen**[♣], **Lilja Øvrelid**[♣], **and Dan Flickinger**[♡]
[♣] University of Oslo, Department of Informatics
[♡] Stanford University, Center for the Study of Language and Information
{angelii|oe|liljao}@ifi.uio.no, danf@stanford.edu

Abstract

We investigate aspects of interoperability between a broad range of common annotation schemes for syntacto-semantic dependencies. With the practical goal of making the LinGO Redwoods Treebank accessible to broader usage, we contrast seven distinct annotation schemes of functor–argument structure, both in terms of syntactic and semantic relations. Drawing examples from a multi-annotated gold standard, we show how abstractly similar information can take quite different forms across frameworks. We further seek to shed light on the representational 'distance' be-

standard data sets (dependency banks) for a range of different languages. These data sets have enabled rigorous evaluation of parsers and have spurred considerable progress in the field of data-driven dependency parsing (McDonald & Nivre, 2011).

Despite widespread use, dependency grammar does not represent a unified grammatical framework and there are large representational differences across communities, frameworks, and languages. Moreover, many of the gold-standard dependency banks were created by automated conversion from pre-existing constituency treebanks—notably the venerable Penn Treebank for English

(Linguistic Annotation Workshop, 2012)

Our Starting Point: Underspecified Logical Forms

LinGO English Resource Grammar (ERG; Flickinger, 2000)

- Broad-coverage HPSG implementation, developed continually since 1993;
- discriminative parse selection, trained on 38,000 mixed-domain sentences;
- parser interface, among other layers: Minimal Recursion Semantics (MRS);
- operators (modals, scopal adverbs, et al.) fixed; quantifiers underspecified.

Our Starting Point: Underspecified Logical Forms

LinGO English Resource Grammar (ERG; Flickinger, 2000)

- Broad-coverage HPSG implementation, developed continually since 1993;
- discriminative parse selection, trained on 38,000 mixed-domain sentences;
- parser interface, among other layers: Minimal Recursion Semantics (MRS);
- operators (modals, scopal adverbs, et al.) fixed; quantifiers underspecified.

$$\langle\, h_1, e_3,\\ \left|\begin{array}{l} h_4\text{:\_every\_q}(\text{ARG0 } x_6, \text{ RSTR } h_7, \text{ BODY } \_\_), h_8\text{:\_linguist\_n\_1}(\text{ARG0 } x_6),\\ h_2\text{:\_have\_v\_1}(\text{ARG0 } e_3, \text{ ARG1 } x_6, \text{ ARG2 } x_9),\\ h_{10}\text{:\_a\_q}(\text{ARG0 } x_9, \text{ RSTR } h_{12}, \text{ BODY } \_\_), h_{13}\text{:\_obsession\_n\_1}(\text{ARG0 } x_9) \end{array}\right|\\ \{\, h_1 =_q h_2,\ h_7 =_q h_8,\ h_{12} =_q h_{13} \,\}\,\rangle$$

Every linguist has an obsession.

Our Starting Point: Underspecified Logical Forms

Some Basic MRS Terminology

- Elementary Predications (EPs);
-
-
-

Grammar (ERG; Flickinger, 2000)

entation, developed continually since 1993;

rained on 38,000 mixed-domain sentences;

ayers: Minimal Recursion Semantics (MRS);

rbs, et al.) fixed; quantifiers underspecified.

$$
\begin{array}{l}
\langle\, h_1,\, e_3, \\
\quad \left| \begin{array}{l}
h_4\text{:\_every\_q(ARG0 } x_6\text{, RSTR } h_7\text{, BODY \_\_)}, \; h_8\text{:\_linguist\_n\_1(ARG0 } x_6\text{)}, \\
h_2\text{:\_have\_v\_1(ARG0 } e_3\text{, ARG1 } x_6\text{, ARG2 } x_9\text{)}, \\
h_{10}\text{:\_a\_q(ARG0 } x_9\text{, RSTR } h_{12}\text{, BODY \_\_)}, \; h_{13}\text{:\_obsession\_n\_1(ARG0 } x_9\text{)}
\end{array} \right| \\
\quad \{\, h_1 =_q h_2,\, h_7 =_q h_8,\, h_{12} =_q h_{13} \,\} \,\rangle
\end{array}
$$

Every linguist has an obsession.

Our Starting Point: Underspecified Logical Forms

Some Basic MRS Terminology

- Elementary Predications (EPs);
- variables: events
-
-

Grammar (ERG; Flickinger, 2000)

entation, developed continually since 1993;

rained on 38,000 mixed-domain sentences;

ayers: Minimal Recursion Semantics (MRS);

rbs, et al.) fixed; quantifiers underspecified.

$$\langle\, h_1, e_3, \left|\begin{array}{l} h_4\text{:\_every\_q}(\text{ARG0 } x_6, \text{ RSTR } h_7, \text{ BODY } \_\_), h_8\text{:\_linguist\_n\_1}(\text{ARG0 } x_6), \\ h_2\text{:\_have\_v\_1}(\text{ARG0 } e_3, \text{ ARG1 } x_6, \text{ ARG2 } x_9), \\ h_{10}\text{:\_a\_q}(\text{ARG0 } x_9, \text{ RSTR } h_{12}, \text{ BODY } \_\_), h_{13}\text{:\_obsession\_n\_1}(\text{ARG0 } x_9) \end{array}\right| \\ \{\, h_1 =_q h_2, h_7 =_q h_8, h_{12} =_q h_{13} \,\} \,\rangle$$

Every linguist has an obsession.

Our Starting Point: Underspecified Logical Forms

Some Basic MRS Terminology

- Elementary Predications (EPs);
- variables: events and instances;
-
-

Grammar (ERG; Flickinger, 2000)

entation, developed continually since 1993;

rained on 38,000 mixed-domain sentences;

ayers: Minimal Recursion Semantics (MRS);

rbs, et al.) fixed; quantifiers underspecified.

$$
\begin{array}{l}
\langle\, h_1,\, e_3, \\
\quad \left| \begin{array}{l}
h_4\text{:\_every\_q}(\text{ARG0 } x_6,\ \text{RSTR } h_7,\ \text{BODY } \_\_),\ h_8\text{:\_linguist\_n\_1}(\text{ARG0 } x_6), \\
h_2\text{:\_have\_v\_1}(\text{ARG0 } e_3,\ \text{ARG1 } x_6,\ \text{ARG2 } x_9), \\
h_{10}\text{:\_a\_q}(\text{ARG0 } x_9,\ \text{RSTR } h_{12},\ \text{BODY } \_\_),\ h_{13}\text{:\_obsession\_n\_1}(\text{ARG0 } x_9)
\end{array} \right| \\
\quad \{\, h_1 =_q h_2,\ h_7 =_q h_8,\ h_{12} =_q h_{13} \,\}\, \rangle
\end{array}
$$

Every linguist has an obsession.

Our Starting Point: Underspecified Logical Forms

Some Basic MRS Terminology

- Elementary Predications (EPs);
- variables: events and instances;
- ‘distinguished’ variables: ARG0;
-

Grammar (ERG; Flickinger, 2000)

entation, developed continually since 1993;

rained on 38,000 mixed-domain sentences;

ayers: Minimal Recursion Semantics (MRS);

rbs, et al.) fixed; quantifiers underspecified.

$$\left\langle\, h_1,\, e_3,\; \left|\begin{array}{l} h_4\text{:\_every\_q}(\text{ARG0 } x_6,\ \text{RSTR } h_7,\ \text{BODY }\_\_),\ h_8\text{:\_linguist\_n\_1}(\text{ARG0 } x_6), \\ h_2\text{:\_have\_v\_1}(\text{ARG0 } e_3,\ \text{ARG1 } x_6,\ \text{ARG2 } x_9), \\ h_{10}\text{:\_a\_q}(\text{ARG0 } x_9,\ \text{RSTR } h_{12},\ \text{BODY }\_\_),\ h_{13}\text{:\_obsession\_n\_1}(\text{ARG0 } x_9) \end{array}\right| \; \{\, h_1 =_q h_2,\ h_7 =_q h_8,\ h_{12} =_q h_{13} \,\} \right\rangle$$

Every linguist has an obsession.

Our Starting Point: Underspecified Logical Forms

Some Basic MRS Terminology

- Elementary Predications (EPs);
- variables: events and instances;
- 'distinguished' variables: ARG0;
- handles

Grammar (ERG; Flickinger, 2000)

entation, developed continually since 1993;

rained on 38,000 mixed-domain sentences;

ayers: Minimal Recursion Semantics (MRS);

rbs, et al.) fixed; quantifiers underspecified.

$$
\left\langle\, h_1,\ e_3,\ \left|\begin{array}{l} h_4\text{:\_every\_q}(\text{ARG0 } x_6,\ \text{RSTR } h_7,\ \text{BODY } \_\_),\ h_8\text{:\_linguist\_n\_1}(\text{ARG0 } x_6), \\ h_2\text{:\_have\_v\_1}(\text{ARG0 } e_3,\ \text{ARG1 } x_6,\ \text{ARG2 } x_9), \\ h_{10}\text{:\_a\_q}(\text{ARG0 } x_9,\ \text{RSTR } h_{12},\ \text{BODY } \_\_),\ h_{13}\text{:\_obsession\_n\_1}(\text{ARG0 } x_9) \end{array}\right|\ \{\, h_1 =_q h_2,\ h_7 =_q h_8,\ h_{12} =_q h_{13} \,\} \right\rangle
$$

Every linguist has an obsession.

Our Starting Point: Underspecified Logical Forms

Some Basic MRS Terminology

- Elementary Predications (EPs);
- variables: events and instances;
- ‘distinguished’ variables: ARG0;
- handles and handle constraints.

Grammar (ERG; Flickinger, 2000)

entation, developed continually since 1993;

rained on 38,000 mixed-domain sentences;

ayers: Minimal Recursion Semantics (MRS);

rbs, et al.) fixed; quantifiers underspecified.

$$\langle\, h_1, e_3,$$

$$\left|\begin{array}{l} h_4\text{:\_every\_q(ARG0 } x_6\text{, RSTR } h_7\text{, BODY \_\_), } h_8\text{:\_linguist\_n\_1(ARG0 } x_6\text{),} \\ h_2\text{:\_have\_v\_1(ARG0 } e_3\text{, ARG1 } x_6\text{, ARG2 } x_9\text{),} \\ h_{10}\text{:\_a\_q(ARG0 } x_9\text{, RSTR } h_{12}\text{, BODY \_\_), } h_{13}\text{:\_obsession\_n\_1(ARG0 } x_9\text{)} \end{array}\right|$$

$$\{\, h_1 =_q h_2,\ h_7 =_q h_8,\ h_{12} =_q h_{13} \,\}\,\rangle$$

Every linguist has an obsession.

From Logical Forms to Dependency Graphs

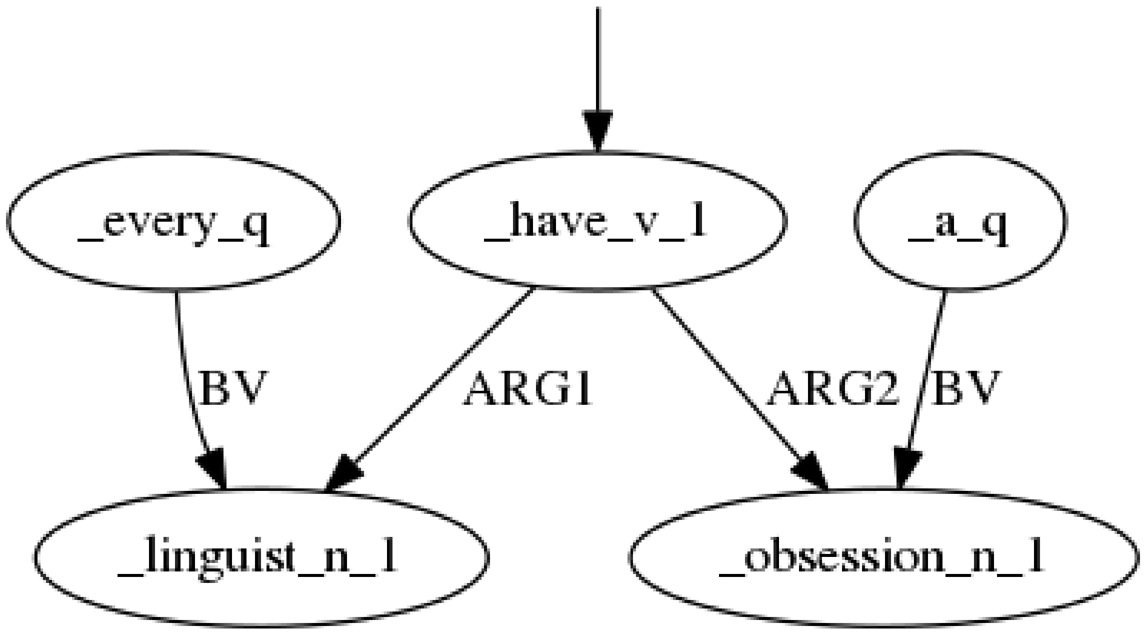
$$\langle\, h_1, e_3, \\ \left| \begin{array}{l} h_4\text{:\_every\_q}(\text{ARG0 } x_6, \text{ RSTR } h_7, \text{ BODY } \_\_), h_8\text{:\_linguist\_n\_1}(\text{ARG0 } x_6), \\ h_2\text{:\_have\_v\_1}(\text{ARG0 } e_3, \text{ ARG1 } x_6, \text{ ARG2 } x_9), \\ h_{10}\text{:\_a\_q}(\text{ARG0 } x_9, \text{ RSTR } h_{12}, \text{ BODY } \_\_), h_{13}\text{:\_obsession\_n\_1}(\text{ARG0 } x_9) \end{array} \right| \\ \{\, h_1 =_q h_2, h_7 =_q h_8, h_{12} =_q h_{13} \,\}\,\rangle$$

EDS: Elementary Depedency Structures (Oepen & Lønning, 2008)

- Eliminate logical variables: one graph node for each elementary predication;
- relation names (predicate constants) become node labels, e.g. _have_v_1;
- node identities determined by *distinguished variable*; except for quantifiers;
- one dependency edge for each instantiated argument role (excluding ARG0);
- dependency target determined by argument variable; view $=_q$ as equations.

Conversion Result for Our Running Example

$$\langle\, h_1, e_3,\ \left|\begin{array}{l} h_4{:}\_\text{every\_q}(\text{ARG0 } x_6,\ \text{RSTR } h_7,\ \text{BODY } \_\_),\ h_8{:}\_\text{linguist\_n\_1}(\text{ARG0 } x_6), \\ h_2{:}\_\text{have\_v\_1}(\text{ARG0 } e_3,\ \text{ARG1 } x_6,\ \text{ARG2 } x_9), \\ h_{10}{:}\_\text{a\_q}(\text{ARG0 } x_9,\ \text{RSTR } h_{12},\ \text{BODY } \_\_),\ h_{13}{:}\_\text{obsession\_n\_1}(\text{ARG0 } x_9) \end{array}\right|\ \{\, h_1 =_q h_2,\ h_7 =_q h_8,\ h_{12} =_q h_{13} \,\}\,\rangle$$

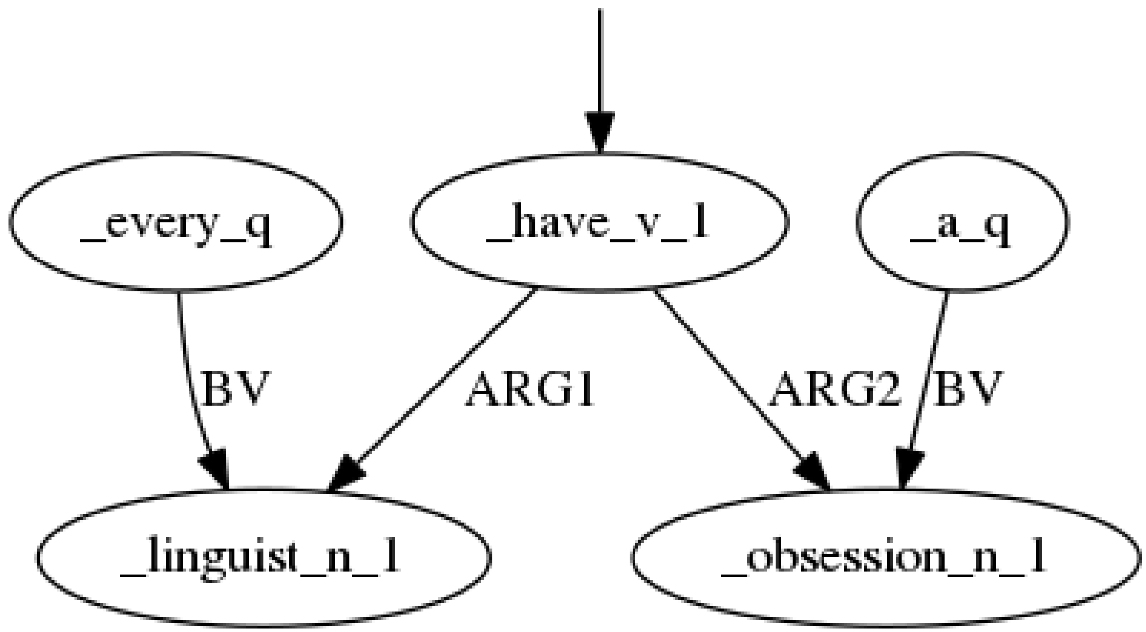


Conversion Result for Our Running Example

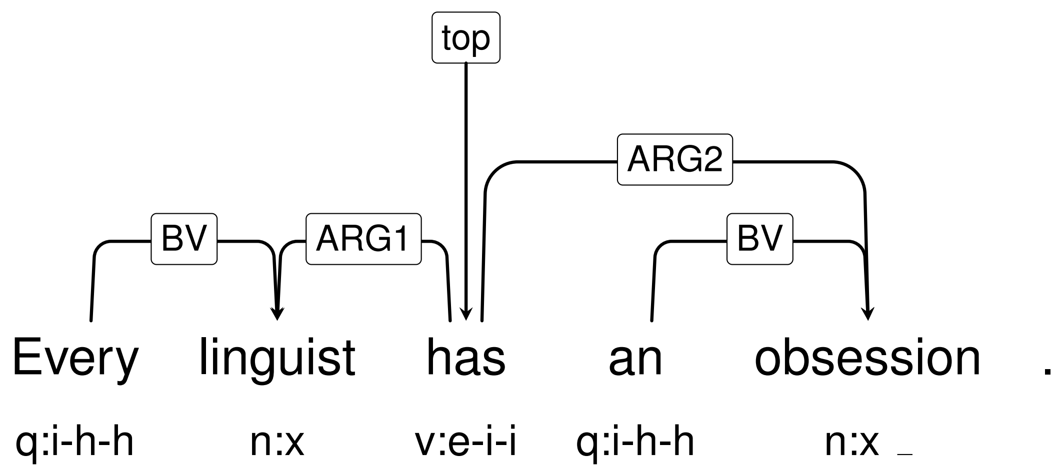
$\langle\, h_1, e_3,$

h_4: every_q(ARG0 x_6, RSTR h_7, BODY), h_8: linguist_n_1(ARG0 x_6),

```
(e3 / _have_v_1
 :ARG1 (x6 / _linguist_n_1 :BV-of (_1 / _every_q))
 :ARG2 (x9 / _obsession_n_1 :BV-of (_2 / _a_q)))
```

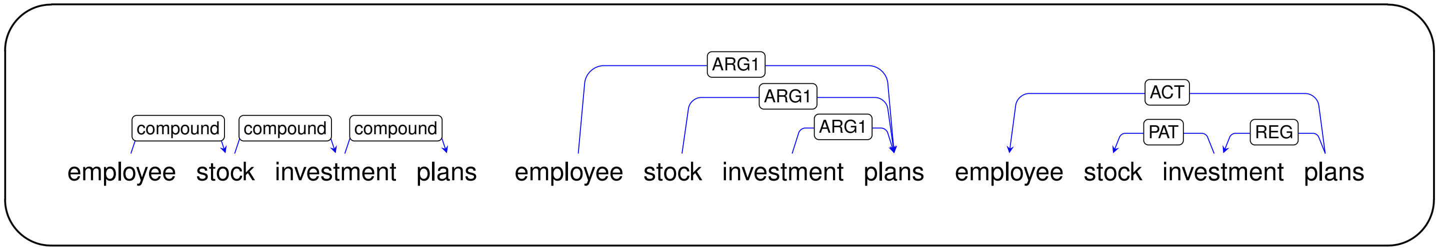


Further Reduction to Bi-Lexical Dependencies



Ivanova et al. (2012): DELPH-IN MRS-Derived Dependencies (DM)

- Lossy reduction of EDS graph: only surface tokens available as nodes;
- (some) construction semantics as edge labels; coarse argument frames;
- argument sharing: graph re-entrancies; vacuous words: unattached nodes;
- designated *top* node (not root): semantic head, highest-scoping predicate.



SDP 2014 (SemEval Task 8)

Broad-Coverage Semantic Dependency Parsing

Stephan Oepen

Universitetet i Oslo & Universität Potsdam

Marco Kuhlmann, Yusuke Miyao, Daniel Zeman,
Dan Flickinger, Jan Hajič, Angelina Ivanova, Yi Zhang

sdp-organizers@emmtee.net

Parallel Annotations of the Venerable WSJ Corpus

DM: DELPH-IN MRS-Derived Bi-Lexical Dependencies

- DeepBank: Fresh HPSG-style annotation, including logical-form semantics;
- ‘lossy’ reduction of MRS meaning representations to bi-lexical dependencies.

Parallel Annotations of the Venerable WSJ Corpus

DM: DELPH-IN MRS-Derived Bi-Lexical Dependencies

- DeepBank: Fresh HPSG-style annotation, including logical-form semantics;
- ‘lossy’ reduction of MRS meaning representations to bi-lexical dependencies.

PAS: Enju Predicate–Argument Structures

- Enju Treebank: Projection of (complete) PTB syntax to HPSG derivations;
- semantic analyses take form of lexicalized predicate–argument structures.

Parallel Annotations of the Venerable WSJ Corpus

DM: DELPH-IN MRS-Derived Bi-Lexical Dependencies

- DeepBank: Fresh HPSG-style annotation, including logical-form semantics;
- 'lossy' reduction of MRS meaning representations to bi-lexical dependencies.

PAS: Enju Predicate–Argument Structures

- Enju Treebank: Projection of (complete) PTB syntax to HPSG derivations;
- semantic analyses take form of lexicalized predicate–argument structures.

PSD: Parts of the Prague Tectogrammatical Layer

- Include all nodes from Prague *t-trees* that correspond to surface tokens;
- re-attach functors of generated nodes; project dependencies to conjuncts.

Parallel Annotations of the Venerable WSJ Corpus

DM: DELPH-IN MRS-Derived Bi-Lexical Dependencies

- DeepBank: Fresh HPSG-style annotation, including logical-form semantics;
- 'lossy' reduction of MRS meaning representations to bi-lexical dependencies.

PAS: Enju Predicate–Argument Structures

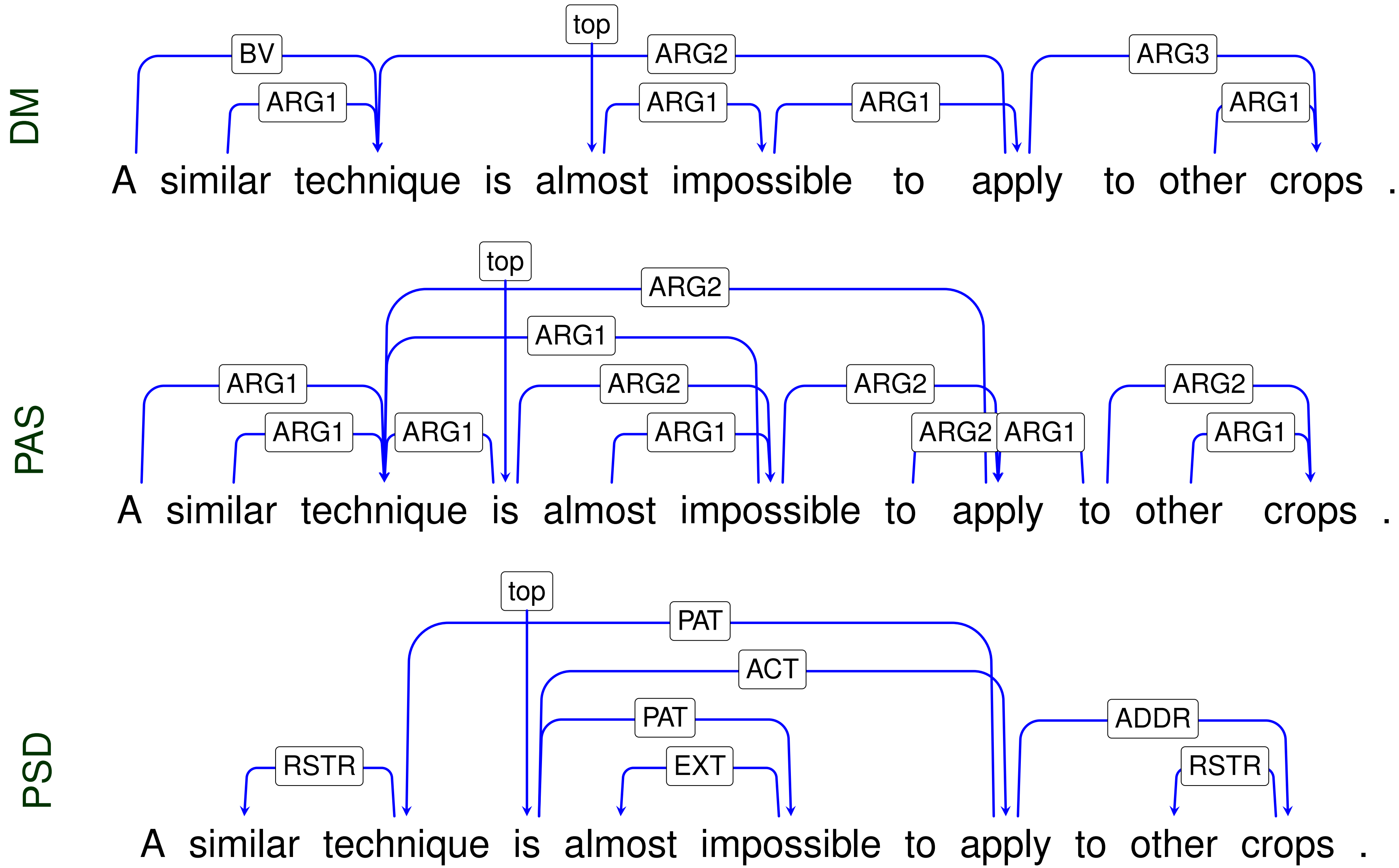
- Enju Treebank: Projection of (complete) PTB syntax to HPSG derivations;
- semantic analyses take form of lexicalized predicate–argument structures.

PSD: Parts of the Prague Tectogrammatical Layer

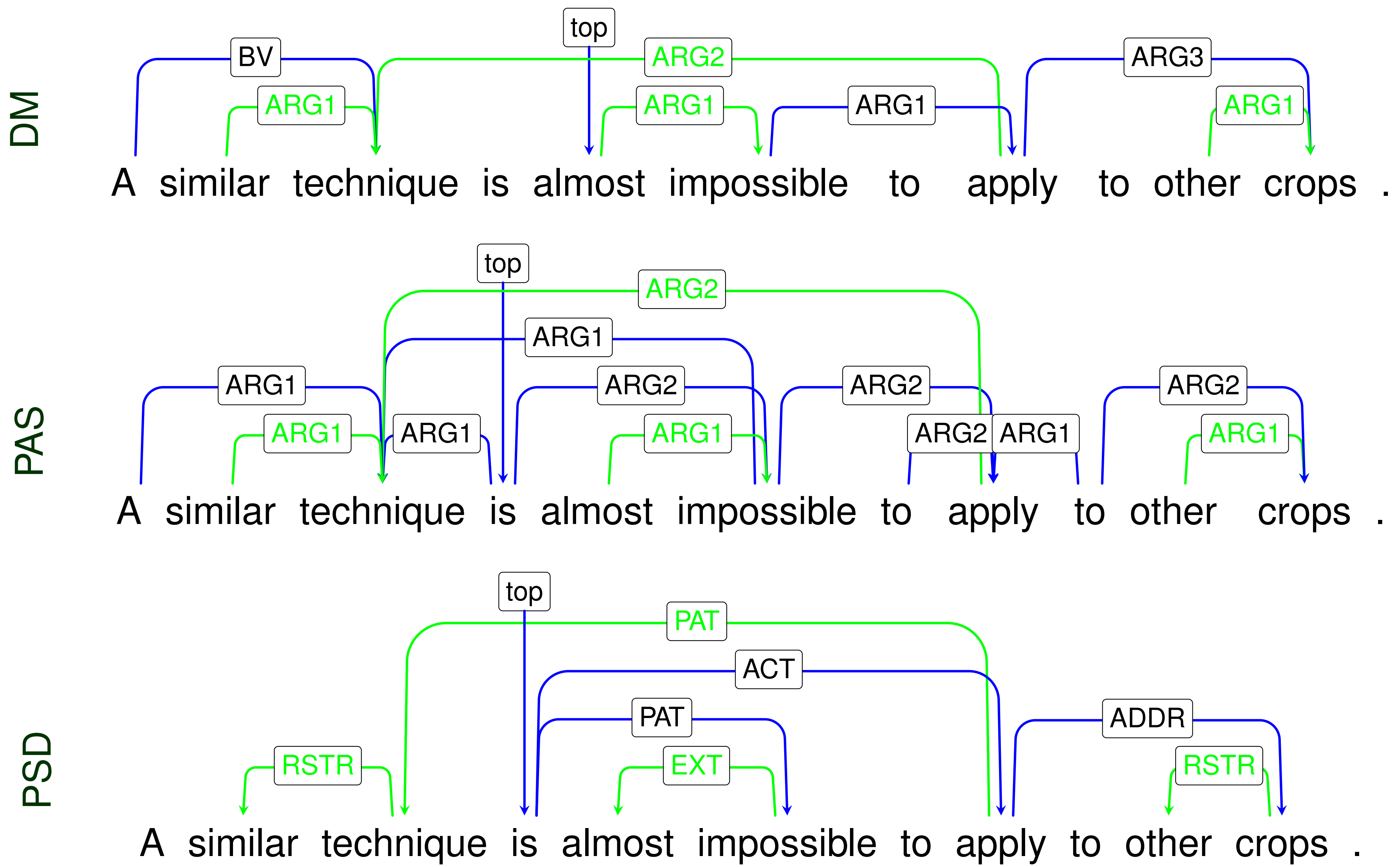
- Include all nodes from Prague *t-trees* that correspond to surface tokens;
- re-attach functors of generated nodes; project dependencies to conjuncts.

Sections 00–20 for Training (802,717 Tokens); Section 21 for Testing (31,948).

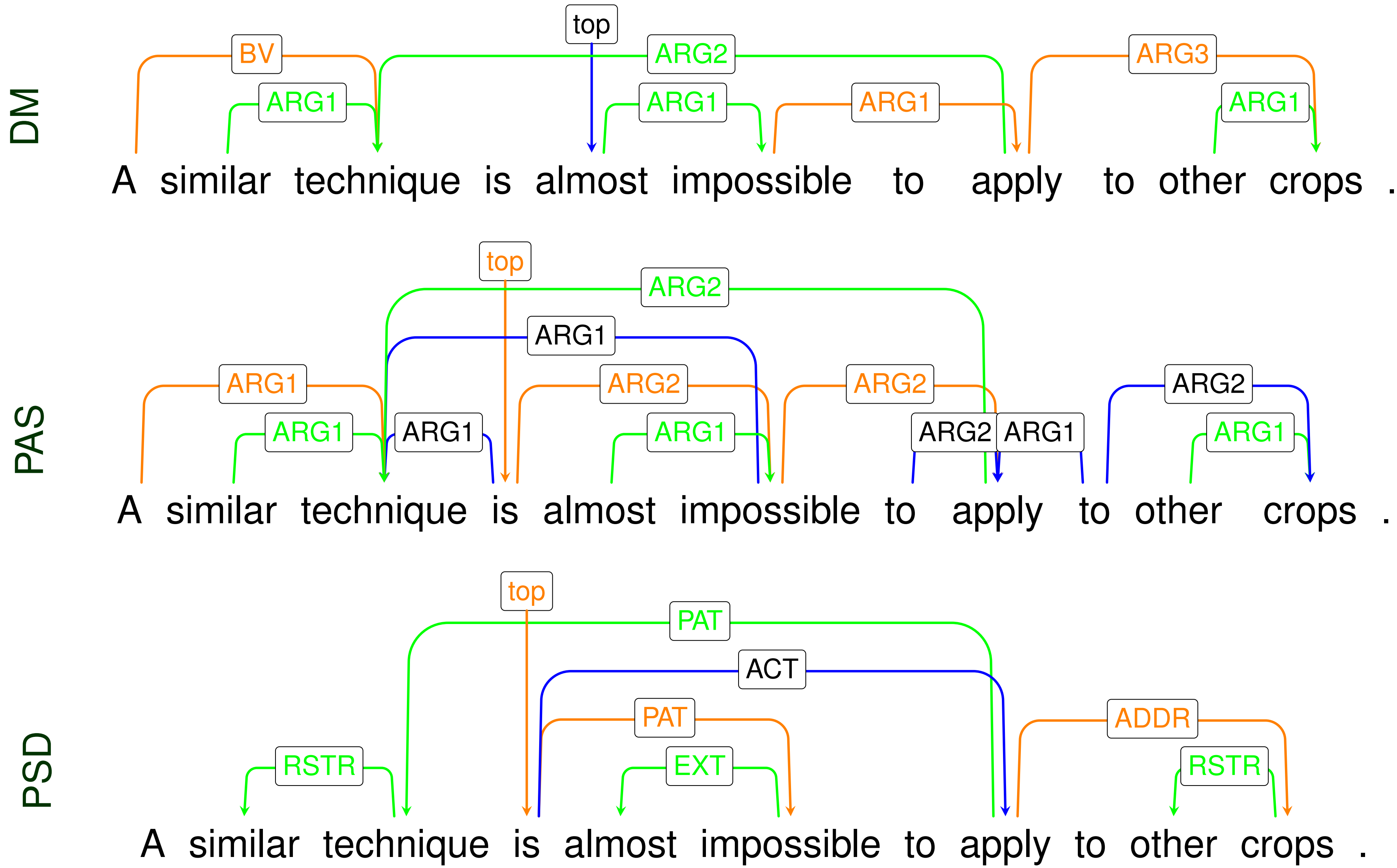
Linguistic Comparison of Target Representations



Linguistic Comparison of Target Representations



Linguistic Comparison of Target Representations



A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42
Copenhagen-Malmö	80.77	84.78	84.04	84.41	20.33	87.69	88.37	88.03	10.16	71.15	68.65	69.88	08.01
Potsdam	77.34	79.36	79.34	79.35	07.57	88.15	81.60	84.75	06.53	69.68	66.25	67.92	05.19
Alpage	76.76	79.42	77.24	78.32	09.72	85.65	82.71	84.16	17.95	70.53	65.28	67.81	06.82
Linköping	72.20	78.54	78.05	78.29	06.08	76.16	75.55	75.85	01.19	60.66	64.35	62.45	04.01

A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42
Copenhagen-Malmö	80.77	84.78	84.04	84.41	20.33	87.69	88.37	88.03	10.16	71.15	68.65	69.88	08.01
Potsdam	77.34	79.36	79.34	79.35	07.57	88.15	81.60	84.75	06.53	69.68	66.25	67.92	05.19
Alpage	76.76	79.42	77.24	78.32	09.72	85.65	82.71	84.16	17.95	70.53	65.28	67.81	06.82
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.01

Observations

- Ensemble system (including graph parsers) best in 'closed' track;

A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42
Copenhagen-Malmö	80.77	84.78	84.04	84.41	20.33	87.69	88.37	88.03	10.16	71.15	68.65	69.88	08.01
Potsdam	77.34	79.36	79.34	79.35	07.57	88.15	81.60	84.75	06.53	69.68	66.25	67.92	05.19
Alpage	76.76	79.42	77.24	78.32	09.72	85.65	82.71	84.16	17.95	70.53	65.28	67.81	06.82
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.01

Observations

- Ensemble system (including graph parsers) best in ‘closed’ track;
- high per-dependency scores: $76 - 92$ F_1 for best ‘closed’ systems;

A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42
Copenhagen-Malmö	80.77	84.78	84.04	84.41	20.33	87.69	88.37	88.03	10.16	71.15	68.65	69.88	08.01
Potsdam	77.34	79.36	79.34	79.35	07.57	88.15	81.60	84.75	06.53	69.68	66.25	67.92	05.19
Alpage	76.76	79.42	77.24	78.32	09.72	85.65	82.71	84.16	17.95	70.53	65.28	67.81	06.82
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.01

Observations

- Ensemble system (including graph parsers) best in ‘closed’ track;
- high per-dependency scores: 76 – 92 F_1 for best ‘closed’ systems;
- exact match sentence accuracy a bit less encouraging: 9 – 38 %;

A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42
Copenhagen-Malmö	80.77	84.78	84.04	84.41	20.33	87.69	88.37	88.03	10.16	71.15	68.65	69.88	08.01
Potsdam	77.34	79.36	79.34	79.35	07.57	88.15	81.60	84.75	06.53	69.68	66.25	67.92	05.19
Alpage	76.76	79.42	77.24	78.32	09.72	85.65	82.71	84.16	17.95	70.53	65.28	67.81	06.82
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.01

Observations

- Ensemble system (including graph parsers) best in 'closed' track;
- high per-dependency scores: 76 – 92 F_1 for best 'closed' systems;
- exact match sentence accuracy a bit less encouraging: 9 – 38 %;
- parsers based on (only) tree approximations not fully competitive;

A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42
Copenhagen-Malmö	80.77	84.78	84.04	84.41	20.33	87.69	88.37	88.03	10.16	71.15	68.65	69.88	08.01
Potsdam	77.34	79.36	79.34	79.35	07.57	88.15	81.60	84.75	06.53	69.68	66.25	67.92	05.19
Alpage	76.76	79.42	77.24	78.32	09.72	85.65	82.71	84.16	17.95	70.53	65.28	67.81	06.82
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.01

Observations

- Ensemble system (including graph parsers) best in ‘closed’ track;
- high per-dependency scores: 76 – 92 F_1 for best ‘closed’ systems;
- exact match sentence accuracy a bit less encouraging: 9 – 38 %;
- parsers based on (only) tree approximations not fully competitive;
- PAS overall easiest to parse, (labeling) PSD is noticeably harder;

A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42

Comparison

- graph adaptation of ('syntactic') TurboParser as best 'open' system;

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Priberam	86.27	90.23	88.11	89.16	26.85	92.56	90.97	91.76	37.83	80.14	75.79	77.90	10.68
CMU	82.42	84.46	83.48	83.97	08.75	90.78	88.51	89.63	26.04	76.81	70.72	73.64	07.12
Turku	80.49	80.94	82.14	81.53	08.23	87.33	87.76	87.54	17.21	72.42	72.37	72.40	06.82
Potsdam	78.60	81.32	80.91	81.11	09.05	89.41	82.61	85.88	07.49	70.35	67.33	68.80	05.42
Alpage	78.54	83.46	79.55	81.46	10.76	87.23	82.82	84.97	15.43	70.98	67.51	69.20	06.60
In-House	75.89	92.58	92.34	92.46	48.07	92.09	92.02	92.06	43.84	40.89	45.67	43.15	00.30

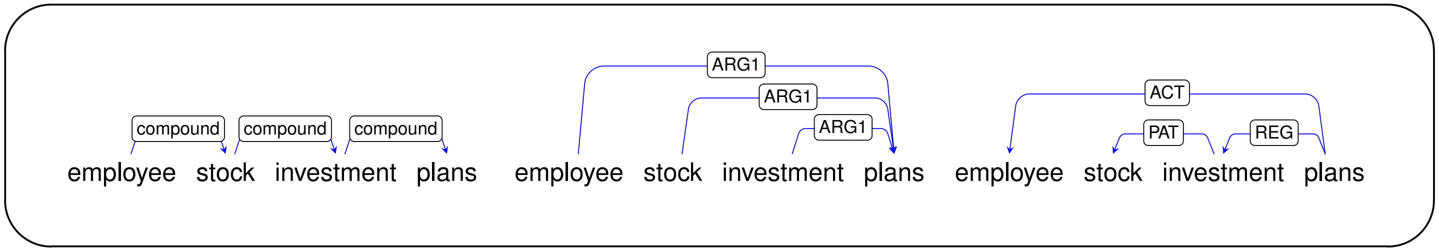
A Glimpse at the SDP State of the Art

		DM				PAS				PSD			
	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Peking	85.91	90.27	88.54	89.40	26.71	93.44	90.69	92.04	38.13	78.75	73.96	76.28	11.05
Priberam	85.24	88.82	87.35	88.08	22.40	91.95	89.92	90.93	32.64	78.80	74.70	76.70	09.42
C													.01
													.19
													.82
													.01

Comparison

- graph adaptation of ('syntactic') TurboParser as best 'open' system;
- (full) 'In-House' systems perform several F_1 points ahead of the field.

	$\overline{\text{LF}}$	LP	LR	LF	LM	LP	LR	LF	LM	LP	LR	LF	LM
Priberam	86.27	90.23	88.11	89.16	26.85	92.56	90.97	91.76	37.83	80.14	75.79	77.90	10.68
CMU	82.42	84.46	83.48	83.97	08.75	90.78	88.51	89.63	26.04	76.81	70.72	73.64	07.12
Turku	80.49	80.94	82.14	81.53	08.23	87.33	87.76	87.54	17.21	72.42	72.37	72.40	06.82
Potsdam	78.60	81.32	80.91	81.11	09.05	89.41	82.61	85.88	07.49	70.35	67.33	68.80	05.42
Alpage	78.54	83.46	79.55	81.46	10.76	87.23	82.82	84.97	15.43	70.98	67.51	69.20	06.60
In-House	75.89	92.58	92.34	92.46	48.07	92.09	92.02	92.06	43.84	40.89	45.67	43.15	00.30



SDP 2016

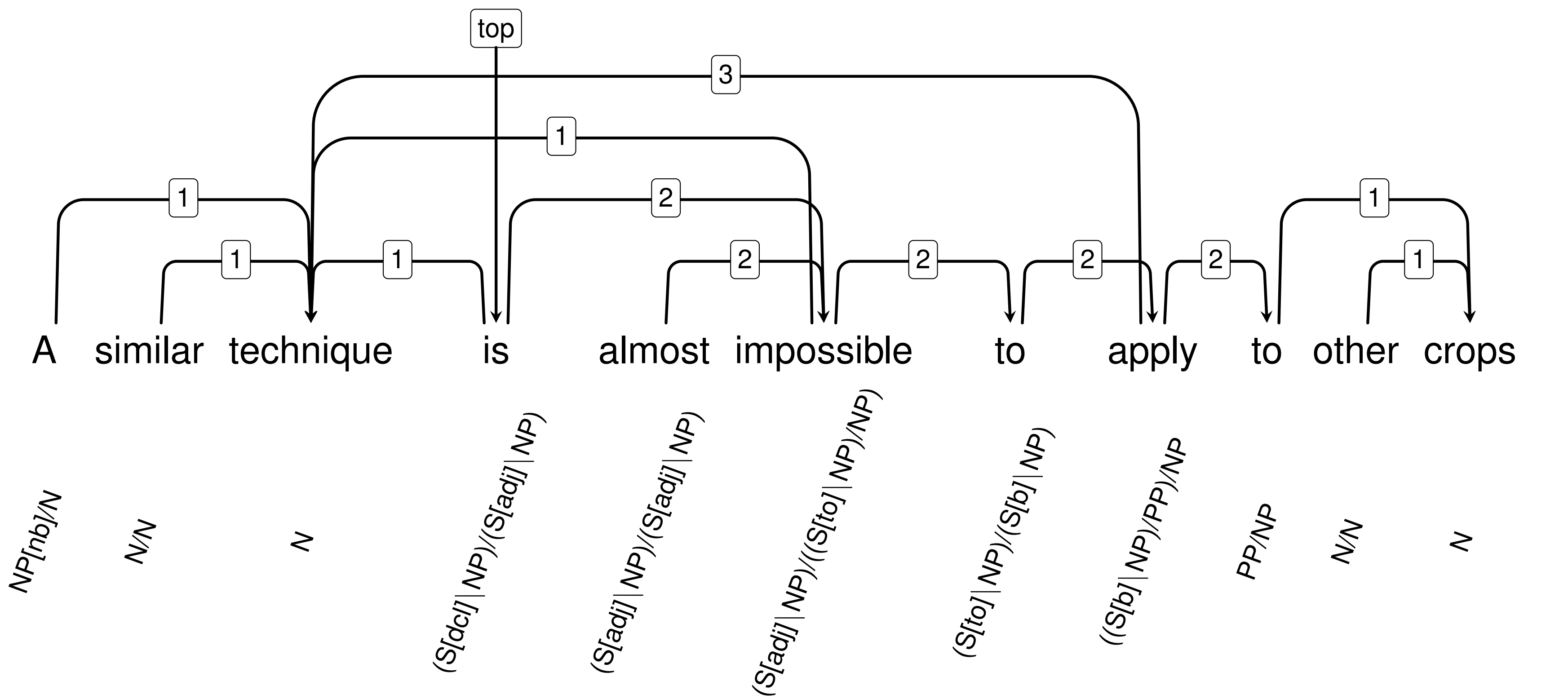
Towards Comparability of Linguistic Graph Banks for Semantic Parsing

Stephan Oepen
Marco Kuhlmann, Yusuke Miyao, Daniel Zeman,
Silvie Cinková, Dan Flickinger, Jan Hajič,
Angelina Ivanova, Zdeňka Urešová

`sdp-organizers@delph-in.net`

New in 2016: CCG Word–Word Dependencies

CCD



CCD: Canonical Conversion from CCGbank

- Connect lexical dependencies with properties from derivation in CCGbank;
- CCG categories as 'frame' identifiers; edge labels identify argument position.

Towards An Algebraic Study of Linguistic Graphs

			CCD	DM	PSD	EDS	AMR	AMR^{-1}
COUNTS	(01)	number of graphs	39604	35656	35656	35656	10309	10309
	(02)	average number of tokens	23.47	22.51	22.51	22.51	20.62	20.62
	(03)	average nodes per token	0.88	0.77	0.64	0.99	0.67	0.67
	(04)	number of edge labels	6	59	90	10	135	100
‘TREENESS’	(05)	$\%_g$ trees	1.45	2.31	42.26	0.98	52.48	18.60
	(07)	average treewidth	1.742	1.303	1.614	1.352	1.524	1.524
	(09)	average edge density	1.070	1.019	1.073	1.047	1.065	1.065
	(10)	$\%_n$ reentrant	28.09	27.43	11.41	28.42	5.23	18.95
	(11)	$\%_g$ cyclic	1.28	0.00	0.00	0.04	3.15	0.71
	(12)	$\%_g$ not connected	12.53	6.57	0.70	1.49	0.00	0.00
	(13)	$\%_g$ multi-rooted	99.67	99.49	99.33	98.75	0.00	77.50
ORDER	(15)	average edge length	2.582	2.684	3.320	–	–	–
	(16)	$\%_g$ noncrossing	48.23	69.21	64.61	–	–	–
	(17)	$\%_g$ pagenumber two	98.64	99.55	98.07	–	–	–

Finally, Some Questions (1/2)

Granularity and Specificity (‘Amount of Information’)

- What (types of) information are lost when going from MRS to EDS to DM?
- How do granularity (and ‘arbitrary’ structural properties) affect parsability?
- How do granularity–structure–parsability trade-offs affect downstream tasks?

Finally, Some Questions (1/2)

Granularity and Specificity (‘Amount of Information’)

- What (types of) information are lost when going from MRS to EDS to DM?
- How do granularity (and ‘arbitrary’ structural properties) affect parsability?
- How do granularity–structure–parsability trade-offs affect downstream tasks?

Dependency Parsing for Norwegian

- Study mapping of NorGram c- and f-structures to bi-lexical dependencies;
- compare linguistically and practically with UD—maybe further conversion?

Finally, Some Questions (2/2)

The Road Ahead in Meaning Representation

- Generation (with the ERG) from different forms of semantic dependencies;
- joint work with AMR folks: Who (and why) needs quantification and scope?
- Cross-fertilization and more uniformity across frameworks and languages.

Finally, Some Questions (2/2)

The Road Ahead in Meaning Representation

- Generation (with the ERG) from different forms of semantic dependencies;
- joint work with AMR folks: Who (and why) needs quantification and scope?
- Cross-fertilization and more uniformity across frameworks and languages.

The Universe and Everything

- The role of grammar: Can we distinguish sentence vs. speaker meaning?
- How can linguistics make a meaningful contribution to NLP, and vice versa?
- Beyond SynSem: Maybe try to put together an MSC Initial Training Network?